Loglinear modelling with latent variables

Luijkx, R.

Published in:
Data analysis

Publication date:
1988

Citation for published version (APA):
Luijkx, R. (1988). Loglinear modelling with latent variables: The case of mobility tables. In W. E. Saris, & I. N. Gallhofer (Eds.), *Data analysis* (pp. 131-159). Macmillan.

7 Loglinear Modelling with Latent Variables: the case of mobility tables*

R. Luijkx

In this chapter mobility tables of Denmark and Britain will be reanalysed. Following Clogg's earlier work (Clogg, 1981a, 1981b), a loglinear analysis with latent variables will be presented. The methodology is embedded in the latent structure analysis, developed by Paul F. Lazarsfeld (Lazarsfeld, 1950a, 1950b, 1954, 1959; Lazarsfeld and Henry, 1968). Besides Lazarsfeld's and Clogg's work, papers by Leo Goodman (1974a, 1974b), Jacques Hagenaars (1976, 1978), Clifford Clogg (1980, 1981a, 1981b) and my own unpublished work (Luijkx, 1983) are the basis of this chapter.

The first part is an exposition of loglinear analysis with latent variables. The latent structure analysis has a clear analogy with factor analysis models for continuous data (Green, 1952). It is a generalisation of the elaboration methodology (Lazarsfeld, 1955; Rosenberg, 1968). I shall restrict the exposition to a two-way contingency table with one latent variable. Because I am dealing with social mobility this is a mobility table. In the past there were problems in estimating the model parameters. Recently developed programs like MLLSA (Clogg, 1977) and LCAG (Hagenaars and van der Walle, 1983) make it possible to compute maximum-likelihood estimates for latent class models in a relatively easy way. LCAG shall be introduced to the reader.

In the second part of this chapter quasi-latent structure models for mobility tables will be presented. Clogg (1981b) points out that latent class models describe the mobility data fairly well and that the parameters of the model are suitable for comparing mobility data of different countries.

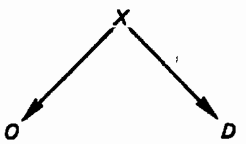
The methodology presented by Clogg (1981b) will be extended.[1] Instead of analysing both mobility tables separately, I shall analyse a three-way table including 'Country' as a variable. This yields an instrument to define constraints on the parameters of the latent structure model across countries. Six models will be presented. These models differ in the equality constraints imposed upon the parameters. The first Model assumes the same model for both countries, but without any constraints on the parameters. There is complete heterogeneity across countries. The last model assumes all the model parameters to be equal for both countries (complete homogeneity). The other four models assume partial homogeneity. Goodness-of-fit statistics for the models will be presented.

7.1 LOGLINEAR MODELLING WITH LATENT VARIABLES

In the elaboration technique the observed relationship among an origin- (O) and a destination- (D) variable is controlled for a third variable (X). Two types of elaboration can be distinguished. In the M(arginal)-type the association among the variables O and D is the same at each level of the (test) variable X ($OD \times X_a = OD \times X_b$), but unequal to the original association ($OD \times X \neq OD$). If $OD \times X$ equals 0 it is called a 'pure type'. In the P(artial)-type the association among the variables O and D is not the same at the levels of X ($OD \times X_a \neq OD \times X_b$). Within the M-type one distinguishes the MA-type, where X is causally prior to O and D and the MI-type where X is an intervening variable between O and D. In the first case X explains the association among O and D; in the second case it interprets the association.

In the latent class analysis, the test variable is not observed: it is a latent variable. In the classical applications the pure MA-type is assumed; the latent variable accounts for the observed relationship among the two variables. Schematically:

Figure 7.1 Basic Latent Class model



Lazarsfeld never succeeded in designing an adequate statistical estimation and testing procedure for latent structure models. He and others used determinantal equations for estimation.[2] It was Leo

Goodman (1974a, 1974b) and Shelby Haberman (1974) who defined algorithms to compute the maximum likelihood estimates (MLEs) of the model parameters. Goodman's algorithm will be discussed.

The following notation will be used:

π	the model parameters
$\hat{\pi}$	the maximum-likelihood estimates of the model parameters
p	the observed sample probabilities
O, D, X	the superscripts for the origin- (O), destination- (D) and latent-variable (X)
i, j, t	the subscripts for, respectively, the classes of the origin-, destination- and latent-variable.

Conditional probabilities are denoted by bars above the superscripts. For example: $\pi_{it}^{\bar{O}X}$ denotes the conditional probability that somebody is in class i with respect to the origin-variable, given that this person is in class t with respect to the latent variable X. The conditional variable is the unbarred variable. Two assumptions are central in the latent class analysis:

- homogeneity: given a class of the latent-variable there is a certain probability for an individual to belong to a class i of the manifest variables. These probabilities are equal for all individuals in a class of a latent-variable. (Individuals can be classified into T mutually exclusive and exhaustive latent classes.)
- local independence: given the latent-variable the manifest variables are conditionally independent.

7.1.1 Unrestricted latent class models

In this example a single latent-variable X is assumed. The observable proportions π_{ij}^{OD} are obtained by collapsing over the levels of latent-variable X. X is only indirectly observed in the table OD. The model is:

$$\pi_{ij}^{OD} = \sum_t \pi_{ijt}^{ODX} \tag{7.1}$$

In Equation (7.1) π_{ijt}^{ODX} gives the probability of an individual being in cell (i, j, t) of the ODX-table.

The law of conditional probability states that O and D are independent in Equation (2). It is X that explains the relationship among O and D. The specification of the local independence assumption is:

$$\pi^{ODX}_{ijt} = \pi^{X}_{t}\, \pi^{\bar{O}X}_{it}\, \pi^{\bar{D}X}_{jt} \tag{7.2}$$

with the (trivial) restrictions:[3]

$$\sum_t \pi^X_t = \sum_i \pi^{\bar{O}X}_{it} = \sum_j \pi^{\bar{D}X}_{jt} = 1 \tag{7.3}$$

Substituting Equation (7.2) into Equation (7.1) gives:

$$\pi^{OD}_{ij} = \sum_t \pi^X_t\, \pi^{\bar{O}X}_{it}\, \pi^{\bar{D}X}_{jt} \tag{7.4}$$

The manifest probabilities π^{OD}_{ij} are the sum of the product of the probability of the latent classses and the conditional probabilities. π^X_t gives the distribution of the classes indirectly observed in the mobility table; $\pi^{\bar{O}X}_{it}$ and $\pi^{\bar{D}X}_{jt}$ describe the distributions of the origin- and destination-variables within each class.

There are several computer programs to compute MLEs of the parameters of latent class models. The first general computer program to my knowledge is Clogg's program MLLSA (Maximum Likelihood Latent Structure Analysis, Clogg, 1977). This program uses Goodman's method of iterative proportional scaling of the estimated parameters (Goodman, 1974b). Haberman's program LAT (Haberman, 1979) uses the scoring algorithm.[4] The first algorithm is also used in LCAG (Latent Class Analysis Goodman: Hagenaars and van der Walle, 1983).[5]

Haberman (1974, pp. 912–15) showed that the maximum-likelihood equations apply for latent structure models in which all the frequency counts are replaced by their estimated expected values given the observed marginals tables.

The program LCAG can be used to estimate parameters of any identifiable hierarchical loglinear model with latent-variables (and also a number of non-hierarchical models). It has some extensions on Goodman's procedure: (1) it is possible to define a hierarchical loglinear model for the relation among latent-variables and (2) one can include latent-variables in models, whose parameters must be estimated according to Goodman's modified-path-analysis approach (Hagernaars and van der Walle, 1983, pp. 2–3). Both features shall be used.[6]

The estimation procedure of the program is based on Goodman (1974b, p. 217, equations 9 to 13b; Hagenaars, 1976):

step 1: start values (initial trial values) for the vector $\hat{\pi}$ of the parameter estimates ($\hat{\pi}^X_t$, $\hat{\pi}^{\bar{O}X}_{it}$, $\hat{\pi}^{\bar{D}X}_{jt}$) are determined (for example, by a random generator).[7]

step 2: given the chosen parameter values the probabilities $\hat{\pi}^{ODX}_{ijt}$ are computed:

$$\hat{\pi}^{ODX}_{ijt} = \hat{\pi}^X_t\, \hat{\pi}^{\bar{O}X}_{it}\, \hat{\pi}^{\bar{D}X}_{jt} \tag{7.5}$$

step 3: the probabilities of the cells in the observed table are computed:

$$\hat{\pi}^{OD}_{ij} = \sum_t \hat{\pi}^{ODX}_{ijt} \tag{7.6}$$

step 4: the conditional probabilities of the observed table given each level of X are computed:

$$\hat{\pi}^{OD\bar{X}}_{ijt} = \hat{\pi}^{ODX}_{ijt} / \hat{\pi}^{OD}_{ij} \tag{7.7}$$

step 5: new estimates of the parameter values are computed. (p^{OD}_{ij} are the observed sample probabilities!)

$$\hat{\pi}^X_t = \sum_{ij} p^{OD}_{ij}\, \hat{\pi}^{OD\bar{X}}_{ijt} \tag{7.8}$$

$$\hat{\pi}^{\bar{O}X}_{it} = [\sum_j p^{OD}_{ij}\, \hat{\pi}^{OD\bar{X}}_{ijt}] / \hat{\pi}^X_t \tag{7.9a}$$

$$\hat{\pi}^{\bar{O}X}_{it} = [\sum_i p^{OD}_{ij}\, \hat{\pi}^{OD\bar{X}}_{ijt}] / \hat{\pi}^X_t \tag{7.9b}$$

After step 5 a next cycle starts at step 2, etc.[8] A latent class is deleted if the estimate tends to zero. This procedure goes on until a convergence-point is reached. There are stop criteria: a maximum number of iterations or a maximum difference in the test-statistics between an iteration and the previous one. After each iteration the parameter are rearranged according to defined retrictions (see the next paragraph) and the sum of the probabilities is rescaled to exactly one.[9]

As test-statistics the Loglikelihood Ratio (L) and the Pearson's *chi*-square (χ^2) are computed with the results of step 3:[10]

$$L = 2n \sum_{ij} p_{ij}^{OD} * \ln (p_{ij}^{OD} / \hat{\pi}_{ij}^{OD}) \tag{7.10a}$$

$$\chi^2 = n \sum_{ij} (p_{ij}^{OD} - \hat{\pi}_{ij}^{OD})^2 / \hat{\pi}_{ij}^{OD} \tag{7.10b}$$

For the unrestricted model, with all parameters identified, the number of degrees of freedom for the *chi*-square statistics is given in the following formula:

$$df = IJ - (I + J - 1)^* T \tag{7.11}$$

It is possible that the obtained solution does not converge on the global maximum, but on a terminal one.[11] The best guarantee to discover which maximum is the maximum-likelihood solution is to repeat the whole procedure with (slightly) different start values. One recognises the maximum-likelihood solution by the lowest value of the test-statistics.[12]

Not every model will yield identifiable parameters. This means that more than one set of parameters can produce the expected frequencies. The estimates of the latent (conditional) probabilities are not uniquely determined by the expected manifest probabilities. One recognises a model with unidentifiable parameters by the fact that different start values give solutions with different parameter estimates, but with an identical test-statistic. It is necessary to impose restrictions upon the parameters to identify the model. But one has to be careful, because some restrictions can make the model unidentifiable (see Goodman, 1974b, pp. 225–6, and the next paragraph). It is better to determine the identifiability of a model beforehand. Goodman (1974b) has presented several problems of identifiability and suggestions to identify models. The parameters of a model are locally identifiable if the transformation defined in Equation (7.2) is non-singular.[13] Both LCAG and MLLSA have a test on (local) identifiability.

7.1.2 Restricted latent class models

I pointed out that it is sometimes necessary to assume certain latent (conditional) probabilities as known to make the parameter estimates

of the model identifiable. There may be other reasons (for example, theoretical ones) to include restrictions.

It is possible to define restrictions of the following kind:

- Assuming known values of the latent class probabilities π_t^x, for example $\pi_2^x = 0.25$ (must be greater than 0 and smaller than 1)
- Assuming known values of the conditional probabilities $\pi_{it}^{\bar{O}X}$ and $\pi_{jt}^{\bar{D}X}$, for example, $\pi_{5\,1}^{\bar{O}X} = 0.20$
- Equality restrictions on the latent class probabilities, for example, $\pi_2^x = \pi_3^x$.
- Equality restrictions on the conditional probabilities, for example, $\pi_{51}^{\bar{O}X} = \pi_{52}^{\bar{O}X}$, $\pi_{51}^{\bar{D}X} = \pi_{41}^{\bar{D}X}$, $\pi_{52}^{\bar{O}X} = \pi_{52}^{\bar{D}X}$, or any other combination.

To determine if the parameter estimates of the restricted latent structure are locally identifiable, a modified form of the evaluation described above is used.[14]

If a restricted model is identified, then the degrees of freedom associated with the model are equal to the number of degrees of freedom for the unrestricted model if that is identified plus the number of non-redundant restrictions imposed upon the model:

$$df = IJ - (I + J - 1)^* T + d \tag{7.12}$$

where d is the number of non-redundant restrictions

As already mentioned above one can also impose restrictions which makes parameter estimates unidentifiable. Consider the following restrictions:

$$\pi_{i1}^{\bar{O}X} = \pi_{i2}^{\bar{O}X} \; (i = 1, \ldots, I) \tag{7.13a}$$

$$\pi_{j1}^{\bar{D}X} = \pi_{j2}^{\bar{D}X} \; (j = 1, \ldots, J) \tag{7.13b}$$

If these restrictions are imposed upon the parameters of the model it is impossible to distinguish between the latent classes *1* and *2*. One must collapse the latent classes *1* and *2*, unless one imposes restrictions on π_j^x or π_2^x.[15] Also, if one only makes the assumptions of Equation (7.13b) the parameter estimates are not identifiable (unless one imposes restrictions upon π_1^x, π_2^x, $\pi_{j1}^{\bar{D}X}$ or $\pi_{j2}^{\bar{D}X}$).[16]

7.2 THE ANALYSIS OF MOBILITY TABLES

Clogg (1981a, 1981b) presents a latent class analysis of mobility tables. He reanalyses two classics among the mobility tables, the British table of Glass (Glass, 1954) and the Danish table of Svalastoga (Svalastoga, 1959). I shall use the same data here. The raw counts are presented in Table 7.1.[17]

Table 7.1 Intergenerational social mobility data of Denmark and of England and Wales

	Son (Denmark)						*Son (England and Wales)*					
	1	*2*	*3*	*4*	*5*		*1*	*2*	*3*	*4*	*5*	
Father 1	18	17	16	4	2	57	50	45	8	18	8	129
2	24	105	109	59	21	318	28	174	84	154	55	495
3	23	84	289	217	95	708	11	78	110	223	96	518
4	8	49	175	348	198	778	14	150	185	714	447	1510
5	6	8	69	201	246	530	0	42	72	320	411	845
	79	263	658	829	562	2391	103	489	459	1429	1017	3497

The conceptual framework of Clogg is based on the model of quasi-independence (Goodman, 1969). Clogg points out that Goodman's quasi-independence model and Hauser's level model (Hauser, 1978, 1979a, 1979b) can, in certain cases, be expressed as (quasi-)latent structure models. I will show to express Goodman's quasi- independence model as latent structure model.

Clogg considers the latent variable as an intervening variable:

Figure 7.2 Latent class model with an intervening variable.

$$O \longrightarrow X \longrightarrow D$$

In this way he defines an MI-type (see above) and he violates Lazarsfeld's rule of defining the latent structure as an MA-type. Theoretically, this specification yields the impression that the latent-variable is explained by the origin-variable and, in its turn, explains the destination-variable. This shows a resemblance with models presented by Boudon (1973). Boudon does not include latent variables in his models, but he defines the variable 'level of education' as an intervening-variable.

Clogg labels the latent variable X as a 'class' variable and the origin- and destination-variable O and D as 'status' variables. The values (classes) of X are to be seen as mixtures of the values (statuses) of O and D. Certain sets of statuses serve to define these classes in probabilistic terms. The pattern of mobility in the OD-table can serve to define classes, within which the structure of mobility takes on a special character, e.g. a class of persons is defined as a group, which possesses random mobility chances with respect to the statuses which together constitute the particular class (Clogg, 1981b, p. 838). According to Clogg O and X can be interpreted as (partly) coexistent, or X can be seen as consequent of O, but $\pi_{it}^{\bar{O}X}$ and $\pi_{jt}^{\bar{D}X}$ cannot be interpreted as simple 'recruitment probabilities' as is normally the case in latent structure models. Clogg defends his view by pointing at the isomorphism he notifies between the latent-variable and Max Weber's class concept (Clogg, 1981a, p. 243, and 1981b, p. 838). But if there is this resemblance this is not a justification (and cannot be one) to define class as an intervening variable.[18] I think Weber's view yields a latent class model in which class causally preceeds both the origin and the destination as in basic latent class model (Figure 7.1).

For the reviewing of the technical aspects of latent class analysis, it is usually not important whether X is defined as either logically preceding or intervening with respect to the O and D variables. The important point is that at least certain classes of O and D define (probabilistically) the classes of X, in which O and D are mutually independent. This is usually not the case if X is a manifest variable. Because this chapter focuses on methodological concepts I shall skip further discussion of Clogg's specification of the relationship among X, O and D and of his theoretical interpretation of the class variable.

7.3 THE TABLES CONSIDERED SEPARATELY

As a baseline model Clogg presents the independence model (perfect mobility), which can be written as follows:

$$\pi_{ij}^{OD} = \pi_t^X \pi_{it}^{\bar{O}X} \pi_{jt}^{\bar{D}X} \tag{7.14}$$

The X variable has only one class and π_t^X is 1 so Equation (7.14) reduces to:

$$\pi_{ij}^{OD} = \pi_i^O \pi_j^D \tag{7.15}$$

In this case there is no latent structure. The mobility would be random, only depending on the marginal distribution of O and D.

Next Clogg assumes two latent classes. The mobility is random within each of the two latent classes, depending only on the marginal distributions $\pi^{\bar{O}X}_{it}$ and $\pi^{\bar{D}X}_{jt}$. In this case there is a 'class barrier'. This model yields unidentifiable parameters.[19] Restrictions have to be imposed upon the parameters. Clogg proposes *a priori* and theoretically founded restrictions. His proposal is to assume some conditional probabilities to be zero. For one class of latent movers the first and for the other class the last status of the origin-variable is assumed to be zero. This states that one latent class of 'movers' cannot consist of members of the first origin status group and another cannot consist of the last (fifth) origin status group. These restrictions are given in (7.16a):[20]

$$\pi^{\bar{O}X}_{11} = \pi^{\bar{O}X}_{52} = 0 \tag{7.16a}$$

$$\pi^{\bar{D}X}_{11} = \pi^{\bar{D}X}_{52}\ 0 \tag{7.16b}$$

It seems logical to impose the same restrictions upon the conditional probabilities of the destination-variable (as in Equation 7.16b). Clogg does not do so on trivial mathematical grounds: because if the same restrictions are defined for both the origin- and destination-variable this yields expected frequencies of zero.

Figure 7.3 Diagrams of the turnover tables (0 is a probability of 0 and 1 is a probability larger than 0)

(1)	(2)	(3)
0 0 0 0 0	1 1 1 1 0	1 1 1 1 0
0 1 1 1 1	1 1 1 1 0	1 1 1 1 1
0 1 1 1 1	1 1 1 1 0	1 1 1 1 1
0 1 1 1 1	1 1 1 1 0	1 1 1 1 1
0 1 1 1 1	0 0 0 0 0	0 1 1 1 1

(1) table for the first latent class of movers
(2) table for the second latent class of movers
(3) total (sum of (1) and (2))

The diagrams in Figure 7.3 have to make this clear. In the first diagram the lower class of 'movers' is displayed (restrictions $\pi^{\bar{O}X}_{11} =$

$\pi^{\bar{D}X}_{11} = 0$). In the second diagram the upper class of 'movers' is displayed (restrictions $\pi^{\bar{O}X}_{52}\ \pi^{\bar{D}X}_{52} = 0$). The result of both restrictions is that the expected frequencies in cells (1, 5) and (5, 1) must always zero: a logical consequence of the combination of restrictions (Equation 7.16a) and (Equations 7.16b). This causes a zero-divide in step 4 of the algorithm.[21]

Clogg also presents a three latent class model with the same and other restrictions, as in the two latent class model.[22] It turns out that none of the three models has a satisfactory fit.

Now we turn to the quasi-latent structure models. Firstly Goodman's quasi-independence model will be described as a quasi-latent structure model. In this case the latent variable has the same number of classes as the origin- and destination- variable plus one. Each cell on the main diagonal is one class of the latent variable. (Cf., for example, Clogg (1980, p. 254).) These classes have the following restrictions:

$$\pi^{\bar{O}X}_{it} = \pi^{\bar{D}X}_{it} = 0 \quad \text{if } i \neq t \quad \text{and} \quad i = 1, \ldots, I \tag{7.17}$$

Because this determines the conditional distributions everybody in class t ($t = 1, \ldots, I$) of the latent variable is a member of the origin and destination class i. Clogg labels these classes of the latent variable deterministic status classes 'latent stayers'. Besides these deterministic status classes one latent class of 'movers' is defined (the last class of the latent variable). No barriers are defined except for the main diagonal.[23] This model is called a quasi-latent structure model because it is not latent in a strict sense: the latent-variable is completely determined by the imposed restrictions.[24] The model reads:

$$\pi^{OD}_{ij} = \sum_t \pi^X_t\, \pi^{\bar{O}X}_{it}\, \pi^{\bar{D}X}_{jt} \quad t = 1, \ldots, 6 \tag{7.18}$$

where $\quad \pi^{\bar{O}X}_{ii}\ \pi^{\bar{D}X}_{ii} = 1$

Clogg defines as a corresponding immobility index:[25]

$$\gamma i = \frac{\pi^X_T\, \pi^{\bar{O}X}_{iT}\, \pi^{\bar{D}X}_{iT} + \pi^X_i}{\pi^X_T\, \pi^{\bar{O}X}_{iT}\, \pi^{\bar{D}X}_{iT}} \tag{7.19}$$

The next model, Clogg defines, is a quasi-latent structure model with a latent variable of $I + 2$ classes. Similar to the former model, I

classes are defined as deterministic status classes (the main diagonal cells). The other two classes are two probabilistic classes of latent 'movers':

$$\pi_{ij}^{OD} = \sum_t \pi_t^x \pi_{it}^{\bar{O}X} \pi_{jt}^{\bar{D}X} \quad t = 1, \ldots, 7 \tag{7.20}$$

where $\quad \pi_{ii}^{\bar{O}X} = \pi_{ii}^{\bar{D}X} = 1$

For these quasi-latent model with two latent classes of 'movers', Clogg defines an immobility index analogous to Equation (7.19). This mobility-ratio measures the surplus of 'stayers' in status i, which cannot be accounted for by the expected immobility under a model with two latent classes of 'movers' (Clogg, 1981b, p. 843). According to Clogg (1981b, p. 860) the problem with earlier immobility ratios was the fact that the observed frequencies in a status group were compared with the expected frequencies given perfect mobility. Whether Clogg's index gives a better interpretation of immobility depends on the fit and the meaning of his model.[26]

Clogg (1981a, 1981b) analyses the British and Danish mobility data separately and compares the differences in the parameter estimates. He does not use a formal test to compare equalities among the parameter estimates in both countries. In the following analysis I will do so by including 'Country' as a variable in the model.

Clogg defines a variant of the quasi-latent structure models.[27] He does not define all deterministic status classes, but only three out of five (the diagonal cells 1, 3 and 5). For purposes of comparison I shall follow Clogg in distinguishing these five latent classes. The first latent class is the upper class and $\pi_{51}^{\bar{O}X}$ is set zero; the second latent class is the lower class with the restriction that $\pi_{12}^{\bar{O}X}$ equals zero. The third through the fifth latent classes correspond with the cells (1, 1), (3, 3) and (5, 5) on the main diagonal. Formally the model is defined as follows:

$$\pi_{ij}^{OD} = \sum_t \pi_t^x \pi_{it}^{\bar{O}X} \pi_{jt}^{\bar{D}X} \quad t = 1, \ldots, 5 \tag{7.21}$$

where $\quad \pi_{13}^{\bar{O}X} = \pi_{34}^{\bar{O}X} = \pi_{55}^{\bar{O}X} = \pi_{13}^{\bar{D}X} = \pi_{34}^{\bar{D}X} = \pi_{55}^{\bar{D}X} = 1$

$\qquad\qquad \pi_{51}^{\bar{O}X} = \pi_{12}^{\bar{O}X} = 0$

In the notation used by Goodman model (Equation 7.21) can be written as $[XO, XD]$. The notation does not take into account the restrictions. It must be remembered that the restrictions mentioned

in Equation (7.21) hold for all of the following models. In all the models the latent variable 'Class' has five classes: two latent classes of 'movers' and three deterministic status classes. Model $[XO, XD]$ has six degrees of freedom. The number of degrees of freedom equals the number of observed frequencies minus one minus the total of independently estimated parameters. The number of parameters to be estimated is 18.[28]

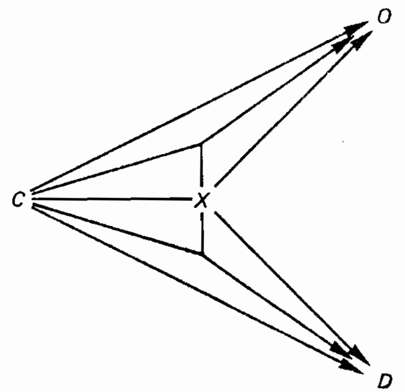
7.4 THE TABLES CONSIDERED SIMULTANEOUSLY

The variable 'Country' (C) is included in the analysis. A three-way observed table ODC is analysed. The latent-variable consists now of ten classes (five for each of the two countries). A latent 'variable' (Y) must be considered as the joint latent variable (C, X), where X is the latent variable X mention in the last section and C is the latent-variable country. The manifest variable C is an accurate indicator of this latent-variable C because of the following restrictions:

$$\pi_{1u}^{\bar{C}Y} = \pi_{2u}^{\bar{C}Y} = 1 \qquad u = 1, \ldots, 5 \quad u = 6, \ldots, 10 \tag{7.22}$$

The result is that the first five levels of the joint variable (C, X) are the latent-variable X for the first country and the second five levels are the latent-variable X for the second country. The model considered now is $[CXO, CXD]$. The test is whether or not the model $[XO, XD]$ holds simultaneously for both countries. In the following the models will be drawn schematically and only the extra restrictions will be mentioned.[29] Model $[CXO, CXD]$:

Figure 7.4 Latent class model [CXO, CXD]



(7.23)

Model $[CXO, CXD]$ is the 'sum' of model $[XO, XD]$ for each country. No extra restrictions are defined, so the number of degrees of freedom equals 2*6 = 12. Although the same model for both countries is assumed, there is complete heterogeneity across countries: both the latent class probabilities and the conditional probabilities are different. The results of the indices of fit are presented in Table 7.2.[30]

Table 7.2 Fit statistics of several models

Model	*df*	*L*	χ^2	Δ	R_L*	*F*
[CXO, CXD]	12	21.24	20.69	1.73	0.99	1.77
[C, X] [CXO, CXD]	16	50.78	50.74	3.02	0.97	3.17
[CX, CO, CD, XO, XD]	20	50.29	47.61	2.57	0.97	2.51
[CX], [CX, CO, CD, XO, XD]	24	94.87	92.25	4.38	0.94	3.95
[CX, XO, XD]	26	216.54	216.36	6.76	0.85	8.33
[C, XO, XD]	30	406.48	404.90	6.57	0.72	13.55

* The baseline model of independence has L = 1465 with df = 32.

Besides χ^2 and L the following measures of fit are reported:

- the dissimilarity index (percentage misclassified: Δ);
- the normed fit index (Bonett and Bentler, 1983, p. 157):

$$R_L = \frac{L_b - L_u}{L_b} \tag{7.24}$$

where L_b is the Loglikelihood Ratio of the restricted model and L_u is the Loglikelihood Ratio of the baseline (relatively unrestricted) model.

- the modified χ^2 (Wonnacott and Wonnacott, 1979, p. 366):

$$F = L/df \tag{7.25}$$

The parameter estimates are (of course) the same as Clogg's (see Table 7.3) and the test-statistic is the sum of the test-statistics for Denmark and Britain separately.[31]

In the next model no association is assumed between the variables C and X. In other words the relative frequency distribution of the

Table 7.3 Parameter estimates for model [CXO, CXD]*

k, t	*prob.*	$\hat{\pi}^X_t$	$\hat{\pi}^{\bar{C}X}_{1t}$	$\hat{\pi}^{\bar{C}X}_{2t}$	$\hat{\pi}^{\bar{O}X}_{1t}$	$\hat{\pi}^{\bar{O}X}_{2t}$	$\hat{\pi}^{\bar{O}X}_{3t}$	$\hat{\pi}^{\bar{O}X}_{4t}$	$\hat{\pi}^{\bar{O}X}_{5t}$	$\hat{\pi}^{\bar{D}X}_{1t}$	$\hat{\pi}^{\bar{D}X}_{2t}$	$\hat{\pi}^{\bar{D}X}_{3t}$	$\hat{\pi}^{\bar{D}X}_{4t}$	$\hat{\pi}^{\bar{D}X}_{5t}$
1,1	0.251	0.102	1.000	0.000	0.071	0.450	0.342	0.137	0.000	0.084	0.384	0.375	0.125	0.032
1,2	0.624	0.254	1.000	0.000	0.000	0.032	0.242	0.466	0.260	0.010	0.022	0.195	0.505	0.269
1,3	0.006	0.003	1.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
1,4	0.059	0.024	1.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
1,5	0.059	0.024	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000
2,1	0.221	0.131	0.000	1.000	0.111	0.507	0.178	0.204	0.000	0.077	0.438	0.165	0.253	0.067
2,2	0.691	0.411	0.000	1.000	0.000	0.043	0.137	0.560	0.260	0.000	0.062	0.117	0.510	0.310
2,3	0.012	0.007	0.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
2,4	0.014	0.008	0.000	1.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
2,5	0.062	0.037	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000

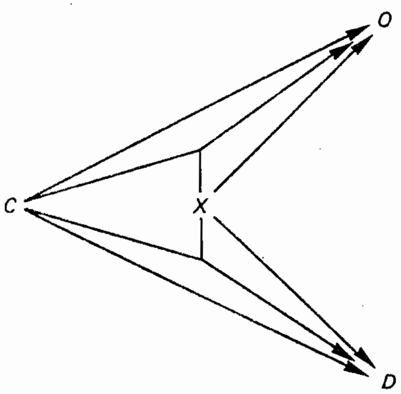
*Because of different sample-size the latent probabilities have to be recalculated to compare them with Clogg's, so the sum of prob. for k=1 and for k=2 adds to one.

latent-variable is the same for both countries. Additional restrictions are:

$$\pi_{t\,1}^{XC} = \pi_{t\,2}^{XC} \; t=1, \ldots, 5 \tag{7.26}$$

This model [C,X][CXO,CXD] has to be considered in the context of the modified-path-analysis approach. Only the relative size of the classes of the latent-variable are assumed equal for both countries. The relative frequency distributions of the origin- and destination-variable given the classes of the latent-variable may differ. This model cannot be straightforwardly estimated because the probabilities $\pi_{t\,k}^{XC}$ are not parameters in the latent class model. Schematically:

Figure 7.5 Latent class model [C,X] [CXO,CDX]



(7.27)

LCAG has, unlike MLLSA, a feature to define loglinear models for any of the marginal tables of (in this case) *CXOD*. Model [C, X] [CXO, CXD] must be considered as the model [C, X] for the marginal table *CX* (table *CXOD* collapsed over *O* and *D*) and [CXO, CXD] for the marginal table *CXOD*. This means *C* and *X* are independent in the table *CX*. As can be seen from the causal scheme in Equation (7.27), *O* and *D* are posterior to both *C* and *X*. To determine the relation among *C* and *X*, *O* and *D* may not be held constant if no null-association between *C* and *X* is assumed. Given this independence of *C* and *X*, the conditional distributions of *O* and *D* given *X* for each country can differ.

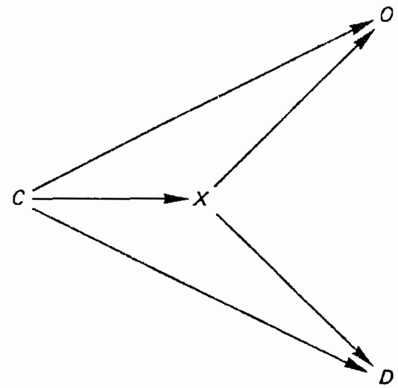
Because model [C, X] [CXO, CXD] has restrictions on the frequency

distribution of the latent variable (the probabilities per class are equal for each country), only four, not 2*4, parameters have to be estimated. The number of degrees of freedom equals 12 + 4 = 16.

The Loglikelihood Ratio is 50.78, which is not a good fit. But for the sake of instruction the parameter estimates are presented in Table 7.4. Both a first glance and residual analysis disclose the reason of this lack of fit: the different levels of inheritance in status group 3.[32]

The next model considered is [CX, CO, CD, XO, XD]. It assumes no higher-order interaction between Country (*C*), Class (*X*), Origin (*O*) and Destination (*D*). The idea is that the distributions of the origin- and destination-variable differ for each country and for the categories of *X*, but not for the cross-classification *CX* (no interactions CXO and CXD). As easy as it is to define the model it is equally difficult to write it as an estimable latent structure model. The strategy is the following. A latent-variable with 250 classes is defined. This is the cross-classification of the latent-variables *C'*, *O'*, *D'* and *X*, the first three of which are perfectly indicated by the manifest variables *C*, *O*, and *D*:

Figure 7.6 Latent class model [CX, CO, CD, XO, XD]



(7.28)

Next a loglinear model is defined among the latent variables *C*, *X*, *O*, *D*.[33] The *chi*-square statistic is 50.29 with 20 degrees of freedom (12 + 2*4 = 20). The results are presented in Table 7.5.[34] It is important to point out here that the conditional probabilities for the two countries differ, but that loglinear parameters (λ and τ) for the relations XO and XD are the same.

In the next model we again assume *C* and *X* to be independent [C, X]

Table 7.4 Parameter estimates for model [C, X] [CXO CXD]*

k, t	prob.	$\hat{\pi}^{x}_{t}$	$\hat{\pi}^{\bar{C}X}_{1t}$	$\hat{\pi}^{CX}_{2t}$	$\hat{\pi}^{OX}_{1t}$	$\hat{\pi}^{OX}_{2t}$	$\hat{\pi}^{\bar{O}X}_{3t}$	$\hat{\pi}^{OX}_{4t}$	$\hat{\pi}^{OX}_{5t}$	$\hat{\pi}^{DX}_{1t}$	$\hat{\pi}^{DX}_{2t}$	$\hat{\pi}^{DX}_{3t}$	$\hat{\pi}^{\bar{D}X}_{4t}$	$\hat{\pi}^{\bar{D}X}_{5t}$
1,1	0.255	0.104	1.000	0.000	0.066	0.408	0.415	0.111	0.000	0.077	0.362	0.447	0.093	0.021
1,2	0.655	0.266	1.000	0.000	0.000	0.042	0.255	0.456	0.247	0.010	0.025	0.210	0.496	0.260
1,3	0.010	0.004	1.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
1,4	0.019	0.008	1.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
1,5	0.061	0.025	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000
2,1	0.255	0.152	0.000	1.000	0.098	0.503	0.177	0.223	0.000	0.068	0.400	0.159	0.286	0.087
2,2	0.655	0.389	0.000	1.000	0.000	0.022	0.132	0.571	0.275	0.000	0.059	0.113	0.511	0.316
2,3	0.010	0.006	0.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
2,4	0.019	0.011	0.000	1.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
2,5	0.061	0.036	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000

*See Table 7.3.

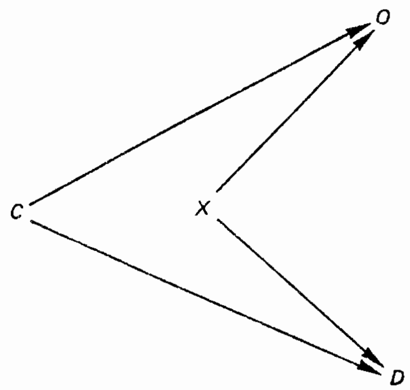
Table 7.5 Parameter estimates for model [CX, CO CD, XO, XO]*

k, t	prob.	$\hat{\pi}^{x}_{t}$	$\hat{\pi}^{CX}_{1t}$	$\hat{\pi}^{\bar{C}X}_{2t}$	$\hat{\pi}^{\bar{O}X}_{1t}$	$\hat{\pi}^{\bar{O}X}_{2t}$	$\hat{\pi}^{\bar{O}X}_{3t}$	$\hat{\pi}^{\bar{O}X}_{4t}$	$\hat{\pi}^{\bar{O}X}_{5t}$	$\hat{\pi}^{DX}_{1t}$	$\hat{\pi}^{\bar{D}X}_{2t}$	$\hat{\pi}^{DX}_{3t}$	$\hat{\pi}^{\bar{D}X}_{4t}$	$\hat{\pi}^{\bar{D}X}_{5t}$
1,1	0.259	0.105	1.000	0.000	0.069	0.434	0.360	0.137	0.000	0.091	0.331	0.351	0.183	0.045
1,2	0.616	0.250	1.000	0.000	0.000	0.034	0.234	0.470	0.262	0.006	0.040	0.204	0.486	0.265
1,3	0.006	0.004	1.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
1,4	0.059	0.024	1.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
1,5	0.060	0.024	1.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000
2,1	0.180	0.107	0.000	1.000	0.135	0.563	0.173	0.129	0.000	0.073	0.511	0.186	0.182	0.047
2,2	0.795	0.433	0.000	1.000	0.000	0.055	0.141	0.561	0.243	0.005	0.065	0.115	0.516	0.299
2,3	0.013	0.007	0.000	1.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000
2,4	0.014	0.008	0.000	1.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000
2,5	0.065	0.038	0.000	1.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.000	1.000

*See Table 7.3.

[CX, CO, CD, XO, XD], and again the modified path analysis approach is used:

Figure 7.7 Latent class model [C,X] [CX,CO,CD,XO,,XD]



(7.29)

The *chi*-square statistic is 94.87 with 24 degrees of freedom (20 + 4 = 24). Again it is cell (3, 3) which causes the bad fit.

The next model is one in which the relative frequency distribution of the latent-variable per country varies, but in which the conditional frequency distributions given the classes of the latent-variable are identical for each country. The following restrictions are defined:[35]

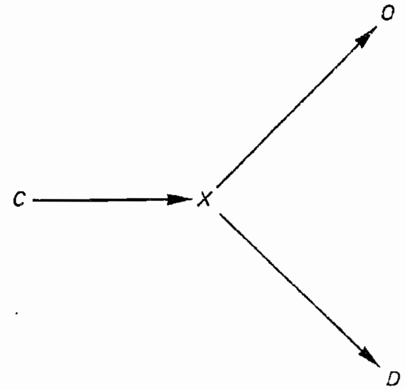
$$\pi_{i\,t\,1}^{OXC} = \pi_{i\,t\,2}^{OXC} \quad i = 1, \ldots, 5 \quad t=1, \ldots, 5 \tag{7.30a}$$

$$\pi_{i\,t\,1}^{DXC} = \pi_{i\,t\,2}^{DXC} \quad i = 1, \ldots, 5 \quad t=1, \ldots, 5 \tag{7.30b}$$

As model [CX, XO, XD] has restrictions on the conditional frequency distribution of the origin- and destination-variables. Fourteen, not 2*14, parameters have to be estimated. The number of degrees of freedom equal 12 + 14 = 26. The model schematically: The fit of this model is 216.54.

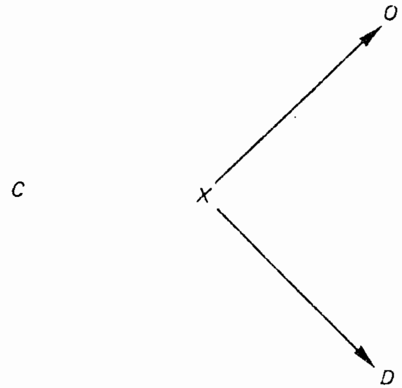
A combination of the restrictions (7.26), (7.30a) and (7.30b) of the models (7.27) and (7.31) yields a model with identical frequency distributions of the latent-variable and identical conditional frequency distributions for the latent-variable for the two countries. Model [C, XO, XD] has both the restrictions of the models [C, X] [CXO, CXD] and [CX, XO, XD] and thus the number of degrees of freedom equals 12 + 4 + 14 = 30. Schematically the model is [C, XO, XD]:

Figure 7.8 Latent class model [CX, XO, XD]



(7.31)

Figure 7.9 Latent class model [C, XO, XD]



(7.32)

The fit of the model is 406.48.[36]

Taking Clogg's analysis as a starting-point, I have shown the possibilities of comparing latent-class models for mobility tables of several countries with an emphasis on loglinear models with latent-variables.

7.5 NOTES

* Direct all correspondence to: Ruud Luijkx, Institute for Sociology, University of Utrecht, Heidelberglaan 2, 3584 CS Utrecht, The Nether-

lands. Earlier versions of this paper were presented at the meeting of the Research Committee on Social Stratification (International Sociological Association) in Budapest on 10 September 1984, and at the First International Conference on Methodological Research (ISA and the Dutch Sociometric Society) in Amsterdam on 5 October 1984. I thank Jacques Hagenaars (Tilburg University, The Netherlands) for his valuable help in the completion of this chapter and Clifford C. Clogg (Pennsylvania State University, Pa, USA) for his comments. Corrections of my English by Kate Ascher are very much appreciated.

1. In a recent article Clogg and Goodman (1984) present a latent structure analysis of a set of contingency tables.
2. This method of determinantal equations is also called the 'Anderson–Lazarsfeld–Dudman method'. It yields consistent estimates only under some conditions, viz. if there are dichotomous variables and less latent classes than $\frac{1}{2}(m+1)$: m being the dimension of the table (see Lazarsfeld and Henry, 1968, chap. 4). This method only give asymptotically efficient parameters in a three-way contingency table with two latent classes (Anderson, 1959). The method can give non-permissible estimates, i.e. not in the interval [0,1] or even complex. See further Anderson (1954), Gibson (1955) and Madansky (1960).
3. Written as a loglinear model:

$$\log m_{ijt}^{ODX} = \mu + \lambda_t^X + \lambda_i^O + \lambda_j^D + \lambda_{it}^{XO} + \lambda_{jt}^{XD} \qquad (7.33)$$

with m the expected frequencies and with the usal restrictions on the parameters.
4. This scoring algorithm is a variant of the Newton–Raphson algorithm (Haberman, 1979, p. 542). The computations resemble those in a weighted regression analysis.
5. Other procedures are (Clogg, 1981a, p. 221): Caroll's canonical decomposition algorithm CANDECOMP (No maximum-likelihood solution), Formann's algorithm based on a logit-type transformation of model parameters, and a certain gradient method.
6. The output of LCAG consists of:

   - the estimated expected probabilities of the classes of the latent-variable;
   - the estimated expected conditional probabilities of the classes of the manifest variables, given the different classes of the latent-variable;
   - the observed and expected frequencies of the cross-classification of the manifest variables;
   - the expected frequencies of the classes of the latent-variables;
   - the *chi*-square statistics L and χ^2.

   It is possible to store the expected frequencies of the cross-classification of the manifest- and latent-variables. These frequencies are essential to compute the loglinear effects (for example, by the use of FREQ, GLIM or ECTA). This makes sense only if the model fits.

7. As initial estimates the results of the determinantal method can be used, if (0,1). Goodman (1974a, pp. 1251–5) also describes how to chose initial trial values in a more sophisticated way.
8. Haberman (1979, pp. 544–7) describes a corresponding iterative algorithm (not the scoring algorithm) in which he determines initial estimates of the frequencies in the observed table, satisfying the loglinear model under study (step 1).
9. The restrictions, assuming parameters to be 0 or 1, can be easily fulfilled by taking the start values of these values as 0 or 1; if other restrictions are assumed, these have to be fulfilled in step 5 of the algorithm at each iteration.
10. It is not easy to decide which test-statistic is the best to use. L, for instance, has nice partionality properties; χ^2, on the other hand, gives a better estimation of the significance-level.
11. Goodman (1974a, p. 1245, note 95) points out that in his experience the iterative procedure did not yield local maxima. Hagenaar's experience is just the opposite (personal communication). A global maximum gives the values of the parameter vector π (π_t^X, π_{it}^{OX}, π_{jt}^{DX}) that maximises the likelihood over the entire range of possible parameter values, where each parameter is strictly within the permissible range [0,1]; a local maximum gives the value of a parameter vector that maximises the likelihood within a neighbourhood of that parameter vector, but not necessarily over the entire range of possible parameter values, where each parameter is strictly within the permissible range; and a terminal maximum gives the value of a parameter vector that maximises the likelihood (globally or locally), but where one or more of the parameters are on the boundary of the permissible range (for example, where some parameter values are either 0 or 1).
12. Clogg also presents a proportional-reduction-in-error measure $\lambda_{X\cdot OD}$ to assess the quality of a latent structure model in terms different from usual criteria based on goodness-of-fit (Clogg, 1981b, p. 840).
13. A short overview of the problem of identifiability is based on Goodman (1974b). The question is whether or not the maximum-likelihood estimates of the parameters in a model are locally identifiable. This depends on how many parameters there are to be estimated. Of the $\hat{\pi}_t^X$ –parameters only $T - 1$ parameters have to be considered; of the $\hat{\pi}_{it}^{OX}$–parameters only $I - 1$ for every t; and of the $\hat{\pi}_{jt}^{DX}$–parameters, only $j - 1$ for every t; see Equation (7.3). So, in total, the number of parameters to be estimated is: $T - 1 + (I - 1)^*T + (J - 1)^*T = (I + J - 1)^*T - 1$.

    The number of expected manifest probabilities equals $IJ - 1$. If Equation (7.34) holds, the number of estimated latent (conditional) probabilities exceeds the number of expected manifest probabilities.

$$IJ - 1 < (I + J - 1)^*T - 1 \qquad (7.34)$$

This is a sufficient condition for unidentifiability. If Equation (7.34) does not hold, one has to calculate the derivative of Equation (7.4) with respect to the expected manifest probabilities. For example:

$$\frac{\partial \hat{\pi}^{OD}_{ij}}{\partial \hat{\pi}^{X}_{t}} = \hat{\pi}^{\bar{O}X}_{it}\,\hat{\pi}^{\bar{D}X}_{jt} - \hat{\pi}^{\bar{O}X}_{iT}\,\hat{\pi}^{\bar{D}X}_{jT}$$

$$\frac{\partial \hat{\pi}^{OD}_{ij}}{\partial \hat{\pi}^{\bar{O}X}_{it}} = \begin{cases} \hat{\pi}^{X}_{t}\,\hat{\pi}^{\bar{D}X}_{jt} & i = 1, \ldots, I-1 \\ -\,\hat{\pi}^{X}_{t}\,\hat{\pi}^{\bar{D}X}_{jt} & i = I \\ 0 & \text{otherwise} \end{cases} \tag{7.35}$$

If the rank of the above-described matrix is equal to the number of columns, the parameter estimates of the model are locally identifiable.

14. In the matrix the columns assumed to have a known value with respect to the parameter estimates are deleted. The columns among which equality constraints are defined with respect to parameter estimates are added. The number of columns of the matrix is now: $(I + J - 1)^*T - 1 + d$, where d is the number of non-redundant restrictions imposed upon the model. Now the rank can be calculated in the same way as described in note 13, above.
15. One can rewrite Equation (7.4) as:

$$\pi^{OD}_{ij} = \sum_{t=2}^{T} \theta^{X}_{t}\,\pi^{\bar{O}X}_{it}\,\pi^{\bar{D}X}_{jt} \tag{7.36}$$

where

$$\theta^{X}_{t} = \begin{cases} \pi^{X}_{l} + \pi^{X}_{2} & t = 2 \\ \pi^{X}_{t} & t = 3, \ldots, T \end{cases}$$

16. One can rewrite (7.4) as:

$$\pi^{OD}_{ij} = \sum_{t=2}^{T} \theta^{X}_{t}\,\theta^{\bar{O}X}_{it}\,\pi^{\bar{D}X}_{jt} \tag{7.37}$$

where θ^{X}_{t} as above

$$\theta^{\bar{O}X}_{it} = \begin{cases} (\pi^{X}_{l}\,\pi^{\bar{O}X}_{il} + \pi^{X}_{2}\,\pi^{\bar{O}X}_{i2}) / \theta^{X}_{2} & t = 2 \\ \pi^{\bar{O}X}_{it} & t = 3, \ldots, T \end{cases}$$

Goodman (1974b, pp. 225, 226) shows that this can be generalised to an m-way table, where the corresponding conditional probabilities in the latent class 1 and 2 are equal for each of the m, resp. m-1 variables.

17. The occupational categories used for the British data are:
    (1) • professional and high administrative
    (2) • managerial and executive

    • inspectional, supervisory and other non-manual (high grade)
    (3) • inspectional, supervisory and other non-manual (low grade)
    (4) • routine grades of non-manual
    • skilled manual
    (5) • semi-skilled manual
    • unskilled manual.

    According to Clogg (1981b, p. 845), Svalastoga collapsed the eight into five categories to make the British and the Danish data comparable.
18. In a recent contribution to class theory, Parkin (1979) looks at (occupational) mobility as class structuring. Schematically:

$$(O - D) \longrightarrow X$$

19. Although the number of parameters to estimate is 17, and thus less than 24 $(IJ-1)$, there exist no immediately visible dependencies.
20. Clogg (1981b, p. 848) points out that if he was not willing to assume that the highest and the lowest status in table 7.1 were known *a priori*, different sets of restrictions could have been used. I would rather have imposed restrictions upon the destination-variable, because it seems more likely that people cannot reach certain positions, given their class, than that they cannot come from certain status groups into certain classes. But, for the purpose of comparison, I shall follow Clogg.
21. LCAG does not stop by a zero-divide, but neglects the step for $\hat{\pi}^{ODX}_{ijt}$. The result is an expected frequency of zero for these cells. In computing the test-statistics, the contribution of the cells with expected frequencies of zero is set to zero. This does not contribute to the Pearson's *chi*-square, but it does to the Loglikelihood Ratio. Therefore the best strategy is to assume the observed frequencies in these cells also to be zero. The same holds true for Haberman's program LAT. Observed frequencies of zero are not a problem in LCAG (because of the arguments just mentioned). But they were in Clogg's case. He used an arbitrary solution by raising the observed frequencies to 0.10 or 1.00. Clogg mentions as another reason the comparability of degrees of freedom among both countries.
22. The total of parameter estimates, without restrictions, would be 26.
23. Mobility across statuses is 'prohibited' for members of each status class (Clogg, 1981b, p. 853).
24. If one of the programs LCAG or MLLSA is used, not all the estimated expected frequencies of the quasi-latent structure model are necessarily identical with those of the quasi-independence model. This follows from the fact that probabilities must be non-zero. So deterministic status classes defined as classes of the latent-variable have a probability of at least zero. Given the model specification it follows logically that only inheritance, and not disinheritance, can be measured. Only a surplus in the main diagonal cells, given the independence model (omitting the diagonal), is measured.
25. This immobility ratio is, as mentioned already (note 24) always greater than 1.
26. There are big problems with these mobility ratios, partly due to the fact

that they are based on truncated equations (see, for example, Hope, 1981; Hauser, 1981).
27. Clogg (1981b, p. 853) points out one can chose any K status classes, where $K \leqq I$.
28. The latent-variable has five classes: the number of parameters to be estimated is 4 (because the sum of the probabilities is 1). The conditional distributions given class 1 and class 2 have five classes for the origin- and destination-variable. The number of parameters to be estimated is five for the origins and five for the destinations per class. Because there are restrictions for the origin classes (one for each class), the number of parameters to be estimated is 3; for the destination classes (no restrictions) this will be 4. The total number of parameters of conditional parameters is 3 + 3 + 4+ 4 = 14. The total number of parameters equals 14 + 4 = 18.
29. If we want to take, in addition to the accurate indicators, also the latent classes into account, the schemata have the following form (I don't define the relations among C', X', O' and D', which differ from model to model):

Figure 7.10 Model Assuming no errors

$$\begin{array}{ccc} & & O' = O \\ C = C' & X & \\ & & D' = D \end{array} \tag{7.38}$$

C', O' and D' being the latent variables and C, O and D being their accurate indicators. But, for the sake of simplicity, we leave the latent-variables out. It must be remembered that the loglinear models are formally defined on the level of the latent-variables. (Cf. LISREL, where the structural model is defined on the level of the latent-variables, even if all are accurately indicated by observed variables.)
30. Clogg (1981b, p. 845) mentions rather complicated sampling arrangements and the impact of the test-statistics on the levels of significance. Therefore these are not reproduced.
31. The slight difference (8.2+12.9) 21.1 v. 21.2 is, in addition to rounding errors, caused by the fact that Clogg replaced the zero entry (5,1) in the data for England and Wales by 0.10.
32. If a modified-path-analysis approach is followed, it can be (as it is in this case) relatively easy but very time-consuming to define a model where there are equalities for all latent classes but the one corresponding with status class 3. I shall introduce these kind of models in the remaining part of this chapter.

33. We can define all the models this way. One reason not to estimate the models this way, if it is not necessary, is the computing time involved.
34. The table looks like 250 latent classes determined by each combination of the cross-classification of the manifest variables C, O and D.
35. Here again we can use the easier way of estimating.
36. Again the bad fit is caused by cell (3,3).

7.6 REFERENCES

ANDERSON, T. W., 'On Estimation of Parameters in Latent Structure Analysis', *Psychometrika*, **19** (1954) 1–10.

—— 'Some scaling models and estimation procedures in the latent class models', in U. Grenander (ed.) *Probability and Statistics* (New York: Wiley, 1959).

BONETT, DOUGLAS C., and BENTLER, P. M., 'Goodness-of-fit Procedures for the Evaluation and Selection of Log-linear Models', *Psychological Bulletin*, **93** (1975) 149–66.

BOUDON, RAYMOND, *Mathematical Structure of Social Mobility* (Amsterdam: Elsevier, 1973).

CLOGG, CLIFFORD C., 'Unrestricted and Restricted Maximum Likelihood Latent Structure Analysis: a manual for users', *Pennsylvania State University, Population Issues Research Office*, Working Paper No. 1977–9 (1977).

—— 'Characterizing the Class Organization of Labor Market Opportunity: a modified latent structure approach', *Sociological Methods & Research*, **8** (1980) 243–72.

—— 'New Developments in Latent Structure Analysis', in D. J. Jackson and E. F. Borgotta (eds) *Factor Analysis and Measurement* (London/Beverly Hills: Sage, 1981a).

—— 'Latent Structure Models of Mobility', *American Journal of Sociology*, **86** (1981b) 836–68.

CLOGG, CLIFFORD C., and GOODMAN, LEO A., 'Latent Structure Analysis of a Set of Multidimensional Contingency Tables', *Journal of the American Statistical Association*, **79** (1984) 762–71.

GIBSON, W. A., 'An extension of Anderson's solution for the latent structure equations' *Psychometrika*, **20** (1955) 69–73.

GLASS, DAVID V. (ed.) *Social Mobility in Britain*, (London: Routledge & Kegan Paul, 1954).

GOODMAN, LEO A., 'On the Measurement of Social Mobility: an index of status persistence', *American Sociological Review*, **34** (1969) 831–50.

—— 'The Analysis of Systems of Qualitative Variables when Some of the Variables are Unobservable. Part I – a modified latent structure approach', *American Journal of Sociology* **79** (1974a) 1179–259.

—— 'Exploratory Latent Structure Analysis using both Identifiable and Unidentifiable Models', *Biometrika*, (1974b) pp. 215–31.

GREEN, BERT F., JR., 'Latent Structure Analysis and its Relation to Factor Analysis', *Journal of the American Statistical Association*, **47** (1952) 71–6.

HABERMAN, SHELBY, 'Log-linear Models for Frequency Tables Derived by Indirect Observation: maximum likelihood equations', *Annals of Statistics*, **2** (1974) 911–24.

—— *Analysis of Qualitative Data*, vol. 2: *New Developments* (New York: Academic Press, 1979).

HAGENAARS, JACQUES A. P., 'Latente Variabelen' (Latent Variables) *Sociale Wetenschappen*, **19** (1976) 250–78.

—— 'Latent Probability Models with Direct Effects between Indicators', *Quality and Quantity*, **12** (1978) 205–22.

HAGENAARS, JACQUES A. P., and VAN DER WALLE, MARTIN A. G., 'LCAG – Een Computerprogramma voor de Analyse van Loglineaire Modellen met Latente Variabelen' (LCAG – A Computerprogram for the Analysis of Loglinear Models with Latent Variables). *Tilburg University: Depart-ment of Sociology* (1983).

HAUSER, ROBERT M. 'A Structural Model of the Mobility Table', *Social Forces*, **56** (1978) 919–53.

—— 'Erratum to HAUSER–1978', *Social Forces*, **57** (1979a) 777.

—— 'Some Exploratory Methods for Modelling Mobility Tables and Other Cross-Classified Data', pp. 413–58 in K. F. Schuessler (ed.) *Sociological Methodology* 1980. (San Francisco: Jossey-Bass, 1979b).

—— 'Hope for the Mobility Ratio', *Social Forces*, **60** (1981) 572–84.

HOPE, KEITH 'The New Mobility Ratio' *Social Forces*, **60** (1981) 544–56.

LAZARSFELD, PAUL FELIX, 'The Logical and Mathematical Foundation of Latent Structure Analysis', chap. 10 in Samuel Stouffer, *Studies in Social Psychology in World War II*, vol. iv: *Measurement and Prediction* (Princeton, N. J.: Princeton University Press, 1950a).

—— 'The Interpretation and Computation of Some Latent Structures', chap. 11 in Samuel Stouffer *Studies in Social Psychology in World War II*, vol. iv: *Measurement and Prediction* (Princeton, N. J.: Princeton University Press, 1950b).

—— 'A Conceptual Introduction to Latent Structure Analysis', pp. 349–87 in P. F. Lazarsfeld (ed.) *Mathematical thinking in the social sciences* (Glencoe, Ill.: The Free Press of Glencoe, 1954).

—— 'Interpretation of Statistical Relations as a Research Operation', pp. 115–25 in P. F. Lazarsfeld and M. Rosenberg (eds) *The Language of Social Research* (Glencoe, Ill.: The Free Press of Glencoe, 1955).

—— 'Latent Structure Analysis', pp. 476–543 in S. Koch (ed.) *Psychology: a study of a science*, vol. 3 (New York: McGraw Hill, 1959 pp. 476–535.

LAZARSFELD, PAUL F., and HENRY, NEIL W., *Latent Structure Analysis* (Boston: Houghton-Mifflin, 1968).

LUIJKX, RUUD, *Een loglineaire analyse van mobiliteitstabellen*. (A loglinear analysis of mobility tables) (Master's thesis, Tilburg University: Department of Sociology, 1983).

MADANSKY, A., 'Determinantal Methods in Latent Class Analysis', *Psychometrika*, **25** (1960) 183–98.

PARKIN, F[illegible] (Cam-
b[illegible]

ROSENBERG, M., *The Logic of Survey Analysis* (New York: Basic Books, 1968).

SVALASTOGA, K., *Prestige, Class and Mobility* (London: Heinemann, 1959).

WONNACOTT, RONALD J., and WONNACOTT, THOMAS H., *Econometrics*, 2nd ed. (New York: Wiley & Son, 1979).